GUARANTEED BOUNDS FOR GENERAL NON-DISCRETE MULTISTAGE RISK-AVERSE STOCHASTIC OPTIMIZATION PROGRAMS

FRANCESCA MAGGIONI* AND GEORG CH. PFLUG[†]

Abstract. In general, multistage stochastic optimization problems are formulated on the basis of continuous distributions describing the uncertainty. Such "infinite" problems are practically impossible to solve as they are formulated and finite tree approximations of the underlying stochastic processes are used as proxies. In this paper, we demonstrate how one can find guaranteed bounds, i.e. finite tree models, for which the optimal values give upper and lower bounds for the optimal values of the original infinite problem. Typically, there is a gap between the lower and upper bound. However, this gap can be made arbitrarily small by making the approximating trees bushier. We consider approximations in the first order stochastic sense, in the convex order sense and based on subgradient approximations. Their use is shown in a multistage risk-averse production problem.

Key words. bounds, barycentric approximations, first order stochastic dominance, convex stochastic dominance, risk measures, multistage stochastic programs

AMS subject classifications. 90C15, 90C90, 65K05

1. Introduction. Multistage stochastic optimization problems are typically formulated in terms of continuous distributions for the underlying stochastic processes describing the problem uncertainty. As such, they are usually intractable as they are originally defined and approximations in terms of finite scenario trees are needed for a numerical solution. The approximation procedure is typically coined as the *scenario tree generation.* Many are the papers dealing with scenario tree generation and the calculation of the approximation error associated to the bounding approach. Even if a large discrete tree model is constructed, the problem might be unsolvable because of the curse of dimensionality. In this situation easy-to-compute bounds have been proposed in literature by solving many much smaller problems instead of the big one associated to the large discrete scenario tree. The idea of looking at sub-problems with only two scenarios goes back to [1] for the two-stage linear case. In [13], approximations of the optimal stochastic solution for multistage linear stochastic programs have been quantified by solving pair sub-problems, by measuring the quality of the deterministic solution and rolling horizon measures which update the estimation and add more information at each stage. In [14] the authors extends the bounding approach of [1, 13, 20], for stochastic multistage mixed integer linear programs, solving a sequence of group sub-problems made by a subset of reference scenarios, and a subset of scenarios from the finite support. Besides, in [15] bounds for multistage convex problems with concave risk functionals as objective are provided. In [3] the authors generalize the definition of the Value of Stochastic Solution and Expected Value of Perfect Information measures for the optimal value of various deterministic equivalent models in a multistage setting with finite support. More recent bounding approaches under the assumption of a given large discrete tree model are proposed in [2, 21, 22]. In [9] the author elaborates an approximation scheme which integrates

---

* Department of Management, Economics and Quantitative Methods, University of Bergamo Via dei Caniana 2, 24127 Bergamo, Italy. Tel: +39-035-2052649 (`francesca.maggioni@unibg.it`).

†University of Vienna. Department of Statistics and Operations Research and International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria. Tel: +43-1-427738630 (`georg.pflug@univie.ac.at`).

stage-aggregation and discretization through coarsening of sigma-algebras to ensure computational tractability, while providing deterministic error bounds.

An alternative approach is to construct not one or a chain of approximating scenario trees, but two trees, a lower tree and an upper tree, the solution of which lead to upper and lower bounds for the optimal value of the original *continuous* problem in terms of the underlying uncertainty. The advantage of the latter approach is that it generates intervals in which the optimal value lies under guarantee. If the size of the guaranteed interval is not satisfactorily small, the two trees may be refined to make the gap smaller.

Results in this direction were for the first time obtained by Frauendorfer [4], followed by [5, 8]. They observe that for convex minimization problems in both the random parameters and decision variables, one has to concentrate the probability mass on the *barycenters* of the partition covering the support of the distribution (a discretization of the underlying probability space), to get a lower bound. For concave functions in the random parameters, one has to distribute the probability mass to the extreme points to get a lower bound. In this way every interior point is a barycenter of a discrete distribution sitting on the extremal points; let us call this method *balayage*[1]. This approach (without naming it balayage) appears in [4, 5]. These authors consider a mixed type of objective, which is concave in one random variable, say η and convex in another one, say ξ and assume they follow a block-diagonal autoregressive process. For upper bounds, one has to use the balayage measure for the convex part and the barycentric measure for the concave part. In contrast, the barycentric measure for the convex part and the balayage measure for the concave part provide a lower bound. The bounds can be made arbitrarily tight by successively partitioning the domain of the random vectors. In [5] the authors discretize the conditional distributions and in a second step the transition probabilities are combined to form a discrete scenario tree. In [8] barycentric discretizations are adopted in a more general setting investigating convex multistage stochastic programs with a generalized non-convex dependence on the random variables.

In this paper, we generalize the bounding ideas of [4, 5, 8] to not necessarily Markovian scenario processes and derive valid lower and upper bounds for the convex case in both the random parameters and decision variables. The concave-convex case requires only an easy additional step, where the upper and lower bound techniques may be used simultaneously for the two parts of the objective. These two parts may be dependent. We construct new discrete probability measures directly from the simulated data of the whole scenario process and not from the discretization of the conditional distributions as done before in the literature.

Our general setup is a multistage stochastic optimization problem of the form

$$v(P) := \min_{x_0,\ldots,x_T} \mathbb{E}[Q(x_0, \xi_1, x_1, \ldots, \xi_T, x_T)] : x_t \lhd \mathcal{F}_t = \sigma(\xi_1, \ldots, \xi_t), x_t \in \mathbb{X}_t\}, \qquad (1.1)$$

where $Q(\cdot)$ is some cost function, $\xi = (\xi_1, \ldots, \xi_T)$ is the stochastic scenario process defined on a probability space $(\Xi, \mathcal{F}, P)$, where $\Xi = \mathsf{X}_{i=1}^T \Xi_i$ and $\mathcal{F} = (\mathcal{F}_1, \ldots, \mathcal{F}_T)$ is the filtration generated by projections of Ξ onto $\mathsf{X}_{i=1}^t \Xi_i$ for each t. The decision process is $x = (x_0, \ldots, x_T)$ and the notation $x_t \lhd \mathcal{F}_t$ means that x_t is measurable w.r.t. to $\mathcal{F}_t$. This constraint is called the nonanticipativity constraint[2]. $\mathbb{X}_t$ is the

---

[1]The word balayage has been introduced by G. Choquet, see for instance [16].

[2]For the sake of simplicity, we do not consider the case where the available information - the filtration - is larger than the one generated by in ξ. In the latter case one has to consider the

set of constraints for the decision variables x_t at stage $t = 1, \ldots, T$, which may be incorporated into the objective function Q by adding the convex indicator function

$$\mathbb{I}_{\mathbb{X}_t}(x_t) = \begin{cases} 0 & \text{if } x_t \in \mathbb{X}_t \\ \infty & \text{otherwise.} \end{cases}$$

In the following, we assume that all constraints have been incorporated in this way into the objective function. $\mathbb{E}$ denotes the mathematical expectation.

Our main problem (1.1) is general in the sense that the scenario process may take (uncountably) infinite values. For the solution of such a problem however, one has to approximate it by a simpler problem, where the scenario process $\tilde{\xi}$ takes only finitely many values. In this case the distribution $\tilde{P}$ of the scenario process $\tilde{\xi}$ is represented by a scenario tree. For the simpler scenario process $\tilde{P}$, the solution of problem (1.1) reduces to an optimization problem in $\mathbb{R}^N$, typically with quite large dimension N. Modern solvers are able to handle it.

Whenever problems are solved by approximation, the question of approximation error arises. Let us mention here a basic approximation result proved in [19].

THEOREM 1.1. *Suppose that $Q(x_0, \xi_1, x_1, \ldots, \xi_T, x_T)$ is convex in all x and Lipschitz with Lipschitz constant L in ξ. Then the approximation error expressed in terms of the absolute difference in optimal values $|v(P) - v(\tilde{P})|$ can be bounded by*

$$|v(P) - v(\tilde{P})| \leq L \cdot d(P, \tilde{P}), \tag{1.2}$$

where $d(P, \tilde{P})$ is the nested distance between the two scenario models P and $\tilde{P}$.

More generally, if the expectation $\mathbb{E}$ in (1.1) is replaced by a distortion functional $\mathcal{R}(Y) = \int_0^1 F_Y^{-1}(u) \, h(u) \, du$, with F_Y being the distribution function of Y, and h being the nonnegative distortion function, then the assertion changes to

$$|v(P) - v(\tilde{P})| \leq L \cdot \sup_{0 \leq u \leq 1} h(u) \cdot d(P, \tilde{P}) \tag{1.3}$$

Remark. The (upper) average value-at-risk (conditional value-at-risk, expected shortfall)

$$AVaR_\alpha(Y) = \min\{q + \frac{1}{1-\alpha} \mathbb{E}[Y - q]_+ : q \in \mathbb{R}\}$$

is a distortion functional with

$$h(u) = \frac{1}{1-\alpha} \mathbb{1}_{\alpha \leq u \leq 1}$$

where $\mathbb{1}$ is the indicator function and thus in (1.3) we have $\sup_{0 \leq u \leq 1} h(u) = (1-\alpha)^{-1}$.

For the exact definition and properties of the nested distance $d(P, \tilde{P})$ we refer to the book [19]. The nested distance is based on minimal transportation costs between the scenario processes ξ and $\tilde{\xi}$. Let us mention here that *any* feasible transportation plan between the two processes leads to an upper bound in (1.2). Thus a possible way of obtaining upper and lower bounds is as follows:

1. Find a finite scenario process $\tilde{\xi}$ with distribution $\tilde{P}$ and a feasible[3] transportation plan π between the infinite process ξ and the finite process $\tilde{\xi}$.

---

nested distribution of ξ (see [18]), but for this paper the multivariate distribution of ξ contains all information.

[3] A transportation plan is feasible, if it respects the filtration structure, see [19].

2. Solve the finite problem and find $v(\tilde{P})$.

3. According to Theorem 1.3 the bounds are then given by:

$$v(\tilde{P}) - L \cdot d_\pi(P, \tilde{P}) \leq v(P) \leq v(\tilde{P}) + L \cdot d_\pi(P, \tilde{P}), \qquad (1.4)$$

where d_π is the distance calculated with the transportation plan π. Obviously the best bounds are obtained for the optimal transportation plan, which defines the nested distance, but any other transportation plan leads also to guaranteed bounds.

However, it turns out, that the bounds given in (1.4) are not very tight: in this paper we propose better ways to find bounds based on the notions of first order and convex dominance for probability measures.

The paper is organized as follows: In section 2 we describe the principles of bounding for two-stage stochastic optimization problems. Section 3 contains the main results for the multistage situation based on first order and convex order stochastic dominance. An example of the construction of upper and lower trees are contained in section 4. Finally Section 5 reports numerical results on a multistage production problem and Section 6 concludes the paper.

2. Bounding two-stage stochastic optimization problems. We consider two-stage stochastic optimization problems where cost function in (1.1) is of the form $Q(x_0, \xi_1, x_1)$ containing just one single random variable $\xi \equiv \xi_1$. We assume here and in the following that the decisions x takes values in $\mathbb{R}^d$, while the random variables ξ take values in $\mathbb{R}^m$.

Let P and P' be probability distributions on $\mathbb{R}^m$. We recall the definition of first order stochastic dominance and convex dominance that will be useful to provide bounds proposed in the paper.

DEFINITION 2.1 (Stochastic Dominance). *Let P and $\tilde{P}$ be probability distributions on $\mathbb{R}^m$. The following stochastic dominances hold true:*

(i) First order stochastic dominance (FSD). *P is dominated by P' in first order sense, and we write*

$$P \prec_{FSD} \tilde{P},$$

if $\int f(v)\, dP(v) \leq \int f(v)\, d\tilde{P}(v)$ for all nondecreasing integrable real valued function f, i.e. for functions f for which $v \leq w$ (componentwise) implies that $f(v) \leq f(w)$.

(ii) Convex stochastic dominance (CXD). *P is dominated by $\bar{P}$ in the convex order sense, and we write*

$$P \prec_{CXD} \bar{P},$$

if $\int f(v)\, dP(v) \leq \int f(v)\, d\bar{P}(v)$ for all convex integrable f.

The relation $\prec_{CXD}$ is also known under the names *Bishop-de Leeuw ordering* or *Lorenz dominance*. More details about order relations can be found in [17]. Typically probabilities being smaller in convex ordering can be obtained by concentrating the mass on the expectation (by virtue of using Jensen's inequality [7]) and probabilities being larger in convex ordering are obtained by moving the masses in a mean-preserving way to the boundaries of the convex hull of the support of P (as is done in Edmundson-Madansky's inequality [11, 12]). These simple facts are the basis of the results of this paper.

2.1. Bounds based on convex order for two-stage stochastic optimization problems.

LEMMA 2.2. *Suppose that the probability measure P can be written as $P = \sum_i w_i P_i$, where w_i are nonnegative weights with $\sum_i w_i = 1$ and P_i are probability measures. Then*

$$\underline{P} := \sum_i w_i \delta_{\mathbb{E}(P_i)} \prec_{CXD} P, \tag{2.1}$$

where $\delta_{\mathbb{E}(P_i)}$ is the point mass associated to $\mathbb{E}(P_i)$ with weight w_i.

Proof. Let g be convex and integrable. Then

$$\int g(v)\,dP(v) = \sum_i w_i \int g(v)dP_i(v) \geq \sum_i w_i g(\mathbb{E}(P_i)) = \int g(v)\,d\underline{P}(v) \ .$$

□

Thus, if the support of P, say Ξ, of the probability P is partitioned into a finite union of disjoint sets $\Xi = \bigcup_i A_i$, P_i be the conditional probability given A_i, i.e. $P_i(B) = P(B \cap A_i)/p_i$ with $p_i = P(A_i)$ and if $z_i = \mathbb{E}(P_i)$, then $\sum_i p_i \delta_{z_i} \prec_{CVX} P$ (see Figure 2.1 for an example).
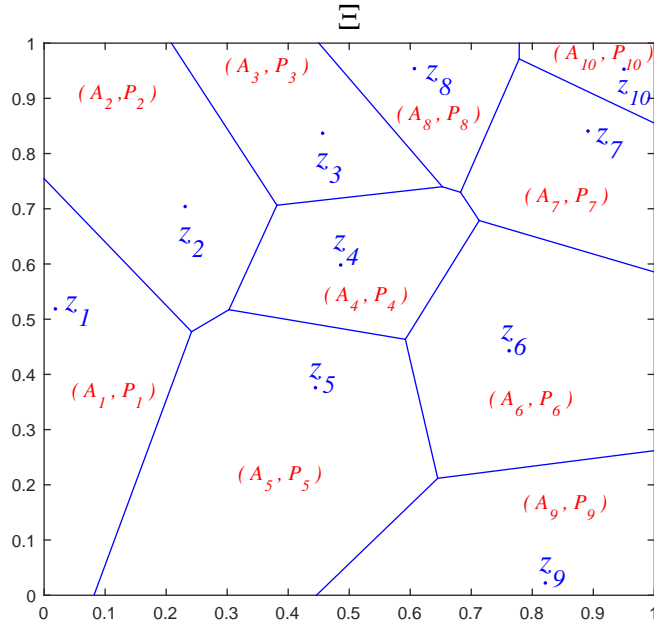

FIG. 2.1. *Example of a partition of a 2-dimensional support Ξ of P, into a finite union of disjoint sets A_i such $\Xi = \bigcup_i A_i$ with expectation $\mathbb{E}(P_i)$.*

To get the inverse relation, suppose that P possesses a Lebesgue density and that the support of P is contained in the union of convex polyhedral sets A_i, such that their interiors are disjoint. A_i is the convex hull of its extremal points, say $\{e_1^i, \ldots, e_{k_i}^i\}$. Each point v in the set A_i it can be written as

$$v = \sum_j w_j^i(v) e_j^i. \tag{2.2}$$

with $w_j^i(v) \geq 0$ and $\sum_j w_j^i(v) = 1$. If the A_i's are simplices, then the representation in (2.2) is unique, but uniqueness is not required here (see Figure 2.2 for an example).
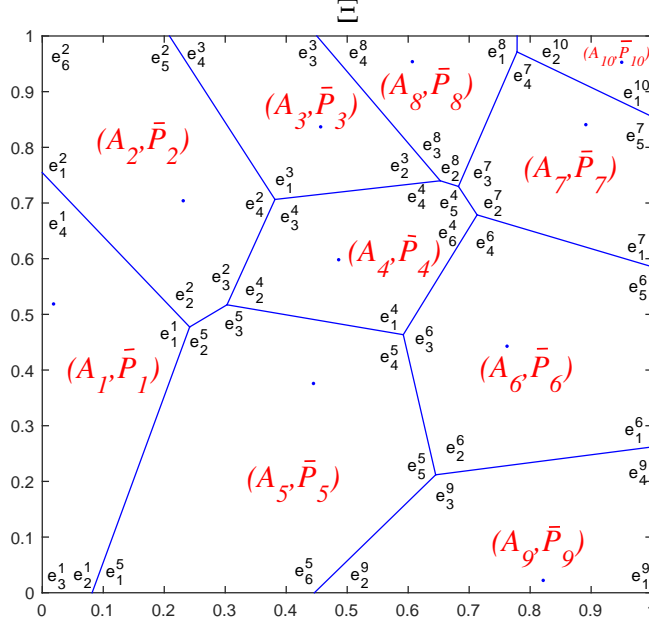

FIG. 2.2. *Example of a partition of a 2-dimensional support Ξ of P, into a finite union of disjoint sets A_i such $\Xi = \bigcup_i A_i$ with extremal points $\{e_1^i, \ldots, e_{k_i}^i\}$ and $\bar{P}_i$ a discrete nonnegative measure sitting on the extremals of A_i.*

LEMMA 2.3. *Let $P_i^v = \sum_j w_j^i(v)\delta_{e_j^i}$ and let $\bar{P}_i = \int_{A_i} P_i^v \, dP(v)$, which is a discrete nonnegative measure sitting on the extremals of A_i. Let $\bar{P} = \sum_i \bar{P}_i$. Then $\bar{P}$ is a probability measure with the property that $P \prec_{CXD} \bar{P}$.*

Proof. Let g be convex and integrable. Then

$$\int g(v) \, d\bar{P}(v) = \sum_i \int_{A_i} g(v) \, d\bar{P}_i = \sum_i \int_{A_i} \sum_j g(e_j^i)w_j^i(v)dP(v)$$

$$\geq \sum_i \int_{A_i} g(\sum_j w_j^i(v)e_j^i) \, dP(v) = \sum_i \int_{A_i} g(v) \, dP(v)$$

$$\geq \int g(v)dP(v).$$

□

Remark: Orderings and risk functionals. If f is nondecreasing and real valued, then $P \prec_{FSD} \tilde{P}$ implies that $P^f \prec_{FSD} \tilde{P}^f$, where P^f resp. $\tilde{P}^f$ are the image measures under f. Thus, if a risk measure ρ is monotonic w.r.t. $\prec_{FSD}$, then $P \prec_{FSD} \tilde{P}$ and $u \mapsto Q(x, u)$ nondecreasing for all x implies that $\rho_P(Q(x, \cdot)) \leq \rho_{\tilde{P}}(Q(x, \cdot))$. Hence also $\inf_x \rho_P(Q(x, \cdot)) \leq \inf_x \rho_{\tilde{P}}(Q(x, \cdot))$.

Likewise, if f is convex and real valued, then $P \prec_{CXD} \bar{P}$ implies that $P^f \prec_{SSD} \bar{P}^f$, where P^f resp. $\bar{P}^f$ are the image measures under f. Thus, if a risk measure ρ is monotonic w.r.t. $\prec_{SSD}$, then $P \prec_{CXD} P'$ and $u \mapsto Q(x, u)$ convex and nondecreasing for all x implies that $\rho_P(Q(x, \cdot)) \leq \rho_{\bar{P}}(Q(x, \cdot))$. Hence also $\inf_x \rho_P(Q(x, \cdot)) \leq \inf_x \rho_{\bar{P}}(Q(x, \cdot))$.

3. Bounding multistage stochastic optimization problems. We consider now a multistage stochastic optimization problem of the form (1.1) where the uncertainty is described by a stochastic process $\xi = (\xi_1, \ldots, \xi_T)$. This process is characterized by P_1, the distribution of ξ_1 and the conditional distributions $\xi_t | (\xi_1 = u_1, \ldots, \xi_{t-1} = u_{t-1})$ for all $t > 1$, denoted by $P_t(\cdot | u_1, \ldots, u_{t-1})$. Our goal is threefold
1. to find tree processes $\bar{P}$ and $\underline{P}$ such that with the value function v as in (1.1) we have

$$v(\underline{P}) \leq v(P) \leq v(\bar{P}),$$

   or with a possible correction term ϵ

$$v(\underline{P}) - \epsilon \leq v(P) \leq v(\bar{P}) + \epsilon;$$

2. to be able to construct refinements of these bounds by considering bushier trees, if the gap is considered too large;
3. to find approximate solutions for the infinite problem (1.1) based in the solutions of the finite problem without solving the infinite problem.

3.1. Bounds based on first order dominance. Stochastic first order dominance given in Definition 2.1 may be broken down in a multistage setting to the conditional distributions. To this end, we introduce the following definition.

DEFINITION 3.1. *We say that a process ξ is totally monotone, if the conditional distributions satisfy*

$$\xi_t | (\xi_1 = u_1, \ldots, \xi_{t-1} = u_{t-1}) \prec_{FSD} \xi_t | (\xi_1 = w_1, \ldots, \xi_{t-1} = w_{t-1})$$

whenever $u_1 \leq w_1, \ldots, u_{t-1} \leq w_{t-1}$ [4].

LEMMA 3.2. *Let the two processes ξ and $\tilde{\xi}$ be totally monotone. Let in addition*

$$\xi_t | (\xi_1 = u_1, \ldots, \xi_{t-1} = u_{t-1}) \prec_{FSD} \tilde{\xi}_t | (\tilde{\xi}_1 = u_1, \ldots, \tilde{\xi}_{t-1} = u_{t-1}). \qquad (3.1)$$

Then $P \prec_{FSD} \tilde{P}$.

Proof. Let f be monotonic in all arguments and let P and $\tilde{P}$ the two probability distributions associated to the totally monotone processes ξ and $\tilde{\xi}$ which satisfy (3.1). Consider

$$f_{T-1}(u_1, \ldots, u_{T-1}) := \int f(u_1, \ldots, u_T) \, dP(u_T | u_1, \ldots, u_{T-1})$$

and

$$\tilde{f}_{T-1}(u_1, \ldots, u_{T-1}) := \int f(u_1, \ldots, u_T) \, d\tilde{P}(u_T | u_1, \ldots, u_{T-1}).$$

---

[4]These are stronger assumptions than introduced in previous papers [6] and [10].

Then f_{T-1} and $\tilde{f}_{T-1}$ are monotonic in $(u_1, \ldots, u_T)$ and by assumption $f_{T-1} \le \tilde{f}_{T-1}$. With the similar argument,

$$f_{T-2}(u_1, \ldots, u_{T-2}) := \int f(u_1, \ldots, u_{T-1}) \, dP(u_{T-1}|u_1, \ldots, u_{T-2})$$

and the analogously defined $\tilde{f}_{T-2}$ are again monotonic and satisfy $f_{T-2} \le \tilde{f}_{T-2}$ and continuing the integration to the end one gets that

$$\int f(\cdot) \, dP(\cdot) \le \int f(\cdot) \, d\tilde{P}(\cdot).$$

☐

If the cost function $Q(x_0, \xi_1, x_1, \ldots, \xi_T, x_T)$ is monotonic in $(\xi_1, \ldots, \xi_T)$, we can apply Lemma (3.2) to problem (1.1) to construct upper bounds.

EXAMPLE 1. *Consider the following multistage stochastic optimization problem with linear constraints in* x.

$$\min \left\{ c_0(x_0) + \mathbb{E}[\sum_{t=1}^{T} c_t(x_t, \xi_t)] \; : \; x \in \mathbb{X} \right\},$$

where the feasible set $\mathbb{X}$ *is given by*

$$
\begin{aligned}
W_0 x_0 &\ge h_0 \\
A_1 x_0 + W_1 x_1 &\ge h_1(\xi_1) \\
A_2 x_1 + W_2 x_2 &\ge h_2(\xi_2) \\
&\vdots \\
A_T x_{T-1} + W_T x_T &\ge h_T(\xi_T) \\
x_1 &\lhd \mathcal{F}_1 \\
&\vdots \\
x_T &\lhd \mathcal{F}_T.
\end{aligned}
\tag{3.2}
$$

Then if $u \mapsto h_t(u)$ *and* $u \mapsto c_t(x_t, u)$ *are monotonically nondecreasing, then the cost function* $Q(x_0, \xi_1, x_1, \ldots, \xi_T, x_T)$ *is monotonic in* $(\xi_1, \ldots, \xi_T)$.

For later use we may also state that if $(x, u) \mapsto c_t(x, u)$ *and* $u \mapsto h_t(u)$ *are convex for all* t, *then the cost function* $Q(x_0, \xi_1, x_1, \ldots, \xi_T, x_T)$ *is jointly convex in all arguments.*

In the following, we elaborate the method for a three stage model, the generalization to more stages can be done analogously, but the notation is more involved. Suppose that (ξ_1, ξ_2) take their values in a rectangle $\Xi = [L_1, U_1] \times [L_2, U_2]$. Let $L_1 = a_1 < a_2 < \cdots < a_{m_a+1} = U_1$ and let for each $i = 1, \ldots, m_a + 1$, $L_2 = b_{i,1} < b_{i,2} < \cdots < b_{i,m_b+1} = U_2$. Define the rectangles $A_{i,j} = [a_i, a_{i+1}] \times [b_{i,j}, b_{i,j+1}]; i = 1, \ldots, m_a, j = 1, \ldots, m_b$. We assume that the scenario distribution (ξ_1, ξ_2) has a density and therefore it does not matter that the rectangles are not disjoint. Define $p_{i,j} = P(A_{i,j})$. Let the finite tree process $\tilde{\xi} = (\tilde{\xi}_1, \tilde{\xi}_2)$ takes in each rectangle $A_{i,j}$ respectively the upper value $\tilde{\xi}_1 = a_{i+1}$ with probability $\sum_j p_{i,j}$ and the right value $\tilde{\xi}_2 = b_{j+1}$ with probability $p_{i,j}/\sum_j p_{i,j}$ conditional on $\tilde{\xi}_1 = a_{i+1}$. Then

$\tilde{\xi} = (\tilde{\xi}_1, \tilde{\xi}_2) := (a_{i+1}, b_{j+1})$ is a finite tree process $\tilde{P}$, for which we solve the basic problem (1.1). Let the solution be $\tilde{x}_1^i$, resp. $\tilde{x}_2^{i,j}$. We extend this to a decision function on Ξ by setting

$$\tilde{x}_1(u_1) = \tilde{x}_1^i \qquad \text{when} \qquad u_1 \in A_i$$

and

$$\tilde{x}_2(u_1, u_2) = \tilde{x}_2^{i,j} \qquad \text{when} \qquad (u_1, u_2) \in A_i \times B_j.$$

We get

$$
\begin{aligned}
v(\tilde{P}) &= \int Q(x_0, u_1, \tilde{x}_1, u_2, \tilde{x}_2) \, d\tilde{P}(u_1, u_2) \\
&= \sum_{i,j} p_{i,j} Q(x_0, a_{i+1}, \tilde{x}_1^i, b_{j+1}, \tilde{x}_2^{i,j}) \\
&\geq \sum_{i,j} \int_{A_i \times B_j} Q(x_0, u_1, \tilde{x}_1^i, u_2, \tilde{x}_2^{i,j}) \, dP(u) \\
&= \sum_{i,j} \int_{A_i \times B_j} Q(x_0, u_1, \tilde{x}_1(u_1), u_2, \tilde{x}_2(u_1, u_2)) \, dP(u) \\
&= \int (Q(x_0, u_1, \tilde{x}_1(u_1), u_2, \tilde{x}_2(u_1, u_2)) \, dP(u) \geq v(P).
\end{aligned}
$$

For establishing a lower bound we use the same setup as before. Let the finite tree process $\underline{\xi} = (\underline{\xi}_1, \underline{\xi}_2)$ takes in each rectangle $A_{i,j}$ respectively the lower value $\underline{\xi}_1 = a_i$ with probability $\sum_j p_{i,j}$ and the left value $\underline{\xi}_2 = b_j$ with probability $p_{i,j} / \sum_j p_{i,j}$ conditional on $\underline{\xi}_1 = a_i$. Then $\underline{\xi} = (\underline{\xi}_1, \underline{\xi}_2) := (a_i, b_j)$ is a finite tree process $\underline{P}$, for which we solve the basic problem (1.1).

If $x_1^*(u_1, u_2)$ and $x_2^*(u_1, u_2)$ are the solutions of the infinite problem, we can make out of then a feasible solution of the finite tree problem, by setting

$$
\begin{aligned}
x_1^+(a_i) &= \min_{u_1 \in A_i} x_1^*(u_1) \\
x_2^+(a_i, b_j) &= \min_{u_1 \in A_i, u_2 \in B_j} x_2^*(u_1, u_2).
\end{aligned}
$$

By monotonicity,

$$
\begin{aligned}
v(P) &= \int Q(x_0, u_1, x_1^*(u_1), u_2, x_2^*(u_1, u_2)) \, dP(u) \\
&\geq \sum_{i,j} \int_{A_{i,j}} Q(x_0, u_1, x_1^+(a_i), u_2, x_2^+(a_i, b_j)) \, dP(u) \geq v(\underline{P}).
\end{aligned}
$$

3.2. Bounds based on convex dominance. In this section we provide bounds based on convex dominance for multistage stochastic programs. One might conjecture that for two scenario processes $P_t(u_t | u_1, \ldots, u_t) \prec_{CXD} \bar{P}_t(u_t | u_1, \ldots, u_t)$ for all t and all $(u_1, \ldots, u_T)$ is sufficient to entail $P \prec_{CXD} \bar{P}$. However this is not true as the following example shows:

EXAMPLE 2. *Let* $\xi_1 \sim N(0, \sigma_1^2)$ *and* $\xi_2 | \xi_1 \sim N(\exp(-\xi_1^2/4), \sigma_1^2)$. *Similarly,* $\bar{\xi}_1 \sim N(0, \sigma_2^2)$ *and* $\bar{\xi}_2 | \bar{\xi}_1 \sim N(\exp(-\bar{\xi}_1^2/4), \sigma_2^2)$. *If* $\sigma_1 < \sigma_2$, *then* $\xi_1 \prec_{CXD} \bar{\xi}_1$ *and also*

$(\xi_2|\xi_1 = x) \prec_{CXD} (\bar{\xi}_2|\bar{\xi}_1 = x)$ *for all* x. *But* $(\xi_1, \xi_2) \not\prec_{CXD} (\bar{\xi}_1, \bar{\xi}_2)$ *as can be seen from the second moments. Choose e.g.* $\sigma_1 = 1/2$, $\sigma_2 = 1$. *Then*

$$\mathbb{E}(\xi_2^2) = \sigma_1^2 + \sqrt{1/\sigma_1^2 + 1/\sigma_1^4} = 4.7221 >$$

$$> 2.4142 = \sigma_2^2 + \sqrt{1/\sigma_2^2 + 1/\sigma_2^4} = \mathbb{E}(\bar{\xi}_2^2).$$

Notice also that a convex order dominating discrete probability cannot by found by choosing dominating discretizations for first components and for all conditional distributions of the second component and concatenating them together as as the next example shows.

EXAMPLE 3. *Suppose that* ξ_1 *is distributed according to Uniform*$[0, 1]$ *and that* $\xi_2|\xi_1$ *is distributed according to Uniform*$[\xi_1(1 - \xi_1), \xi_1(1 - \xi_1) + 1]$. *Let* $\bar{\xi}_1$ *take the values 0 and 1 each with probability 1/2. Then* $\bar{\xi}_1$ *dominates* ξ_1 *in convex order. Likewise, let for each* u, $\bar{\xi}_2(u)$ *take the values* $u(1 - u)$ *and* $u(1 - u) + 1$ *with probability 1/2 each. Then the conditional distributions* $\xi_2|\xi_1 = u$ *are dominated by* $\bar{\xi}_2(u)$ *for all* u. *But if concatenating* $\bar{\xi}_1$ *with the conditional distributions* $\bar{\xi}_2(u)$ *only the conditional distributions for* $u = 0$ *and* $u = 1$ *are used and one obtains that* $(\bar{\xi}_1, \bar{\xi}_2)$ *has a distribution, which sits on all 4 edges of the unit square with equal probabilities 1/4. But this is not a convex dominant of* (ξ_1, ξ_2), *since*

$$\mathbb{E}(\xi_2^2) = 16/30 > \mathbb{E}(\bar{\xi}_2^2) = 1/2.$$

The same example shows that concatenating lower bound approximations for the conditional probabilities does not lead to a lower bound approximation for the total probability. That is why in our procedure, the full multistage sample is used to generate the bounding trees and not just the conditional distributions.

According to the two previous examples we claim that convex order of conditional distributions is not sufficient, and a stronger condition must hold: The construction of convex upper and lower bounds must be based on comparing the whole distributions and not just the conditional ones.

For simplicity, we consider here three-stage stochastic programs. Generalization to more stages requires a more complicated notation but can be done in an analogous way.

Let $Q(x_0, \xi_1, x_1(\xi_1), \xi_2, x_2(\xi_1, \xi_2))$ be a convex function in ξ and let P be a probability measure on a bounded rectangle in $\mathbb{R}^2$. Notice that the extension to $\mathbb{R}^m \times \mathbb{R}^m$ is obvious, but the notation gets more complicated so we omit it. As in (1.1) $\mathcal{F}_1$ is the σ-algebra generated by the first component in $\mathbb{R}^2$ and $\mathcal{F}$ is Borel σ-algebra of $\mathbb{R}^2$. Our problem is reduces to find

$$v(P) = \min_{x_0, x_1, x_2} \mathbb{E}_P[Q(x_0, \xi_1, x_1(\xi_1), \xi_2, x_2(\xi_1, \xi_2))] \tag{3.3}$$

where x_0 is deterministic, x_1 is measurable w.r.t. $\mathcal{F}_1$ and x_2 is measurable w.r.t. $\mathcal{F}$.

3.2.1. Upper tree approximation based on convex order stochastic dominance. Upper bounds for minimization problems are always easy to obtain, since every feasible solution constitutes an upper bound. If the basic problem P contains continuous distributions, but the approximating problem $\tilde{P}$ is discrete, then one has to construct a feasible solution for P out one for $\tilde{P}$. Suppose that $\tilde{P}$ is a scenario tree with values z_1^i in the second stage and $z_2^{i,j}$ in the third stage and $\tilde{x}_1^i$ and

$\tilde{x}_2^{i,j}$ are its discrete solutions of the problem with $\tilde{P}$ as the distribution of the scenario process. Then by any reasonable extension function one may construct a solution for the P-problem (3.3), for instance by setting

$$\tilde{x}_1(\xi_1) = x_1^i \qquad \text{if } z_1^i \text{ is the point, which is closest to } \xi_1$$
$$\tilde{x}_2(\xi_1, \xi_2) = x_2^{i,j} \qquad \text{if } (z_1^i, z_2^{i,j}) \text{ is the point, which is closest to } (\xi_1, \xi_2).$$

Obviously,

$$\min_{x_0, x_1(\cdot), x_2(\cdot)} \mathbb{E}_P[Q(x_0, \xi_1, x_1(\xi_1), \xi_2, x_2(\xi_1, \xi_2))] \leq \mathbb{E}_{\tilde{P}}[Q(x_0, \xi_1, \tilde{x}_1(\xi_1), \xi_2, \tilde{x}_2(\xi_1, \xi_2))].$$

That is any extension of a solution of any tree process $\tilde{P}$ leads to an upper bound. However notice that one has to evaluate the objective function for the scenario process P and the solution $\tilde{x}$ in order to get the upper bound.

We aim however at finding an upper bound, which can be calculated on a finite tree without evaluating the continuous problem. A construction similar to the one for $\bar{P}$ in the two-stage case may be used.

Suppose that (ξ_1, ξ_2) take their values in a rectangle $\Xi = [L_1, U_1] \times [L_2, U_2]$. Let $L_1 = a_1 < a_2 < \cdots < a_{m_a+1} = U_1$ and let for each i, $l_2 = b_{i,1} < b_{i,2} < \cdots < b_{i,m_b+1} = u_2$.

Define the rectangles $A_{i,j} = [a_i, a_{i+1}] \times [b_{i,j}, b_{i,j+1}], i = 1, \ldots, m_a, j = 1, \ldots, m_b$. We assume that the scenario distribution (ξ_1, ξ_2) has a density and therefore it does not matter that the rectangles are not disjoint.

Let $P_{i,j}$ be the distribution P conditioned on the set $A_{i,j}$, i.e.

$$P_{i,j}(D) = \frac{1}{p_{i,j}} P(A_{i,j} \cap D)$$

with $p_{i,j} = P(A_{i,j})$. We have that

$$P = \sum_{i,j} p_{i,j} P_{i,j}.$$

Let $\bar{P}_{i,j}$ be a probability measure sitting on the four extremals of $A_{i,j}$ such that $P_{i,j} \prec_{CXD} \bar{P}_{i,j}$. Then $P \prec_{CXD} \bar{P}$ by the following lemma.

LEMMA 3.3. *If* $P_{i,j} \prec_{CXD} \bar{P}_{i,j}$ *for all* i,j *and* $P = \sum_{i,j} p_{i,j} P_{i,j}$ *and* $\bar{P} = \sum_{i,j} p_{i,j} \bar{P}_{i,j}$ *then*

$$P \prec_{CXD} \bar{P}.$$

Proof. Let f be convex, then

$$\mathbb{E}_P(f) = \sum_{i,j} p_{i,j} \mathbb{E}_{P_{i,j}}(f) \leq \sum_{i,j} p_{i,j} \mathbb{E}_{\bar{P}_{i,j}}(f) = \mathbb{E}_{\bar{P}}(f).$$

$\bar{P}$ is a tree process and one may find the solution of the pertaining optimization problem. Let x_1^{*i} and $x_2^{*i,j}$ be the solution of this problem. We construct a continuous extension of this solution. If $(\xi_1, \xi_2) = \sum w^{i,j}(\xi_1, \xi_2) e_1^{i,j}$, then we set $\bar{x}_1(\xi_1) = \sum w^{i,j}(\xi_1, \xi_2) x_1^{*i}$ and $\bar{x}_2(\xi_1, \xi_2) = \sum \sum w^{i,j}(\xi_1, \xi_2) x_2^{*i,j}$. By construction

$$\mathbb{E}_{\bar{P}}[Q(x_0, \xi_1, \bar{x}_1(\xi_1), \xi_2, \bar{x}_2(\xi_1, \xi_2))] \geq \mathbb{E}_P[Q(x_0, \xi_1, x_1(\xi_1), \xi_2, x_2(\xi_1, \xi_2))].$$

If now the P-problem is solved to optimality, one sees that the relationship

$$v(\bar{P}) \geq v(P)$$

holds. □

3.2.2. Lower tree approximation based on convex order stochastic dominance.
The problem of finding lower bounds is slightly more involved than finding the upper bounds. We refer to the same construction of rectangles as before. Let $z^{i,j}$ be the barycenters of $P_{i,j}$.

By construction, $\underline{P}_{i,j} \prec_{CXD} P_{i,j}$ and by the Lemma 3.3 $\underline{P} \prec_{CXD} P$.

Let $\mathcal{F}^A$ be the σ-algebra generated by the sets $A_{i,j}$. Notice that for all integrable function f

$$\mathbb{E}[f(\xi_1, \xi_2)|\mathcal{F}^A] = \sum_{i,j} \mathbb{E}_{P_{i,j}}(f)\mathbb{1}_{A_{i,j}}.$$

We now distinguish two cases:
- Case I: Conditionally on A_i, ξ_2 is independent of ξ_1.
- Case NI: The independence assumption is not satisfied.

[Case I]. In this case, the conditional expectations given $A_{i,j}$ of all functions which only depend on the first component ξ_1 do not depend on j, in particular $z_1^{i,j} = z_1^i$ does not depend on j.

The following proposition holds true:

PROPOSITION 3.4. *Let $\underline{P}$ be the tree constructed using z_1^i and $z_2^{i,j}$ with scenario probabilities $p_{i,j}$. Then: $v(\underline{P}) \leq v(P)$.*

Proof. We get

$$\begin{aligned}
&\mathbb{E}_P[Q(x_0, \xi_1, x_1(\xi_1), \xi_2, x_2(\xi_1, \xi_2))] \\
&= \mathbb{E}_P\{\mathbb{E}_P[Q(x_0, \xi_1, x_1(\xi_1), \xi_2, x_2(\xi_1, \xi_2))|\mathcal{F}^A]\} \\
&\geq \mathbb{E}_P[Q(x_0, \mathbb{E}(\xi_1|\mathcal{F}^A), \mathbb{E}(x_1|\mathcal{F}^A), \mathbb{E}(\xi_2|\mathcal{F}^A), \mathbb{E}(x_2|\mathcal{F}^A)] \\
&= \mathbb{E}_{\underline{P}}[Q(x_0, z_1^i, x_1^i, z_2^{i,j}, x_2^{i,j})] \\
&\geq v(\underline{P})
\end{aligned}$$

The last inequality comes from the fact that we have constructed a feasible solution for $\underline{P}$, but this need not be the optimal one. However its optimal solution leads a further lower bound for the problem with P as scenario process. Therefore $\underline{P}$ is a lower tree approximation. □

[Case NI]. If the independence assumption is not valid, we use an different approach:

Notice that if $z \mapsto Q(x, z)$ is convex and finitely valued, then for every point $\tilde{z}$

$$Q(x, z) \geq Q(x, \bar{z}) + \langle q(\bar{z}|x), z - \tilde{z} \rangle$$

where $q(\bar{z}|x)$ is a subgradient of $z \mapsto Q(x, z)$ at $\bar{z}$ for fixed x. Consequently

$$\begin{aligned}
\mathbb{E}[Q(x, \xi)] &\geq \mathbb{E}[Q(x, \bar{\xi})] + \mathbb{E}\langle q(\bar{\xi}|x), (\xi - \bar{\xi}) \rangle \\
&\geq \mathbb{E}[Q(x, \bar{\xi})] - [\mathbb{E}(C^2(\bar{\xi}))]^{1/2} \cdot [\mathbb{E}(\|\bar{\xi} - \xi\|^2)]^{1/2}
\end{aligned} \tag{3.4}$$

where $C(\bar{\xi})$ is a bound for the norm of the subgradients $q(\bar{\xi}|x)$ uniformly in x

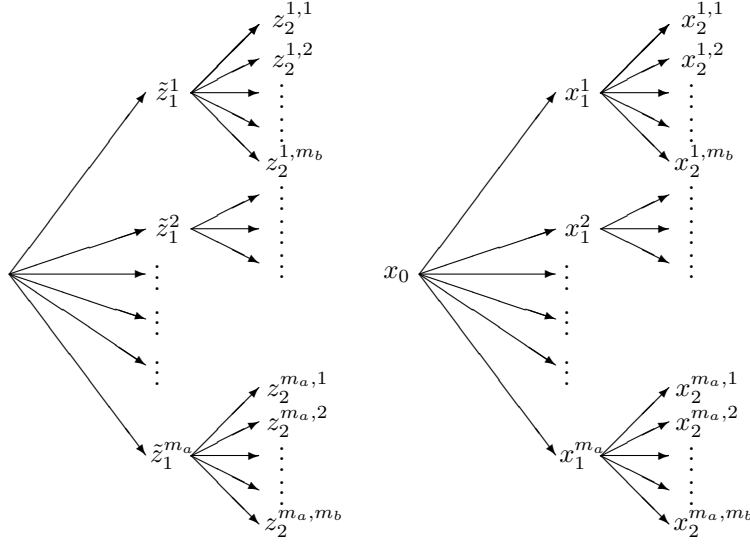$$\sup_x \|q(\bar{\xi}|x)\| \leq C(\bar{\xi}).$$

FIG. 3.1. *Left: The notation of the constructed scenario tree. Right: The pertaining decision tree.*

The case of constraints of the form $(x, z) \in B$ needs a little more attention. Suppose that $Q(x, z) = Q_0(x, z) + \mathbb{I}_B(x, z)$, where $Q_0(x, z)$ is convex and finitely valued and B is a convex set.

If $(x, \bar{z}) \in B$ and $q(x, \bar{z})$ is a subgradient of $z \mapsto Q(x, z)$ at $\bar{z}$ for fixed x, then $Q(x, z) \geq Q(x, \bar{z}) + \langle q(\bar{z}|x), z - \bar{z} \rangle$ even if $(x, z) \notin B$. Therefore (3.4) is a valid lower bound for all values of ξ.

A similar argument holds for the three-stage case: if $z_1 \mapsto Q(x_0, z_1, x_1, z_2, x_2)$ is convex, with subgradient $q(z_1|x_0, x_1, x_2, z_2)$, then

$$\mathbb{E}_P[Q(x_0, \xi_1, x_1, \xi_2, x_2)]$$
$$\geq \mathbb{E}[Q(x_0, \bar{\xi}_1, x_1, \xi_2, x_2)] + \mathbb{E}[\langle q(\bar{\xi}_1|x_0, x_1, x_2, \xi_2), (\xi_1 - \bar{\xi}_1) \rangle]$$
$$\geq \mathbb{E}[Q(x_0, \bar{\xi}_1, x_1, \xi_2, x_2)] - [\mathbb{E}_P(C^2(\bar{\xi}_1))]^{1/2} \cdot [\mathbb{E}_P(\|\xi_1 - \bar{\xi}_1\|^2)]^{1/2}.$$

For the construction of the lower approximating tree, let $z^{i,j} = (z_1^{i,j}, z_2^{i,j})$ as before be the barycenters of $P_{i,j}$.

However, since the lower approximation has to be a tree, we set

$$\bar{z}_1^i = \sum_{j=1}^{m_b} p_{i,j} z_2^{i,j}.$$

Let $\underline{P}$ be the tree constructed using $\bar{z}_1^i$ as first stage values and and $z_2^{i,j}$ as second stage values with scenario probabilities $p_{i,j}$. The notation of this tree as well as of the decision tree is shown in Figure 3.1.

PROPOSITION 3.5. *Let $\underline{P}$ be the tree constructed using $\bar{z}_1^i$ as first stage values and and $z_2^{i,j}$ as second stage values with scenario probabilities $p_{i,j}$. For the tree process $\underline{P}$ we have that*

$$v(\underline{P}) - [\sum_i p_i C(\tilde{z}_1^i)] \cdot \max_{i,j} \|\bar{z}_1^i - z_2^{i,j}\| \leq v(P).$$
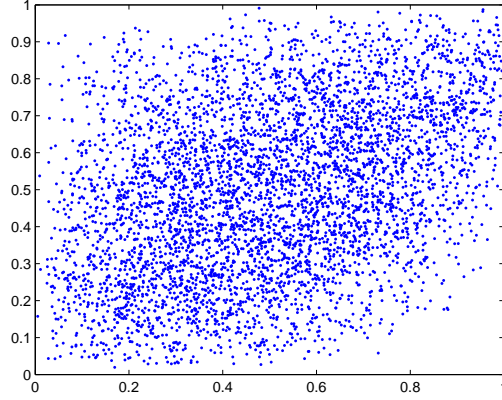
FIG. 4.1. *5000 points distributed according to the distribution P of the scenario process (ξ_1, ξ_2), with $\xi_1 \sim P_1$ distributed according to a Beta(2,2) distribution and $\xi_2|\xi_1 \sim P_2(\cdot|\xi_1)$ according to a Beta$(2, \frac{1.4 - 0.8 \cdot \xi_1}{0.3 + 0.4 \cdot \xi_1})$.*

Proof. As in Proposition 3.4 before, we have that

$$\mathbb{E}_P[Q(x_0, \xi_1, x_1(\xi_1), \xi_2, x_2(\xi_1, \xi_2))] \geq \mathbb{E}_{\underline{P}}[Q(x_0, z_1^{i,j}, x_1^i, z_2^{i,j}, x_2^{i,j})]$$
$$= \mathbb{E}_{\underline{P}}[Q(x_0, \bar{z}_1^i, x_1^i, z_2^{i,j}, x_2^{i,j})] + \mathbb{E}_{\underline{P}}[Q(x_0, z_1^{i,j}, x_1^i, z_2^{i,j}, x_2^{i,j})] - \mathbb{E}_{\underline{P}}[Q(x_0, \bar{z}_1^i, x_1^i, z_2^{i,j}, x_2^{i,j})]$$
$$\geq \mathbb{E}_{\underline{P}}[Q(x_0, \bar{z}_1^i, x_1^i, z_2^{i,j}, x_2^{i,j})] - [\sum_i p_i \, C^2(\bar{z}_1^i)]^{1/2} \cdot [\sum_{i,j} p_{i,j} \|\bar{z}_1^i - z_1^{i,j}\|^2]^{1/2}$$

If the $z_1^{i,j}$ do not coincide for different j but fixed i, then the correction term

$$[\sum_i p_i \, C^2(\bar{z}_1^i)]^{1/2} \cdot [\sum_{i,j} \|\bar{z}_1^i - z_1^{i,j}\|^2]^{1/2}$$

has to be subtracted. Otherwise the correction term disappears. □

4. Lower and upper scenario trees construction: an example. In order to demonstrate the approach proposed in Section 3 with a simple example, assume that distribution P of the scenario process is given by (ξ_1, ξ_2), where $\xi_1 \sim P_1$ is distributed according to a Beta(2,2) distribution and $\xi_2|\xi_1 \sim P_2(\cdot|\xi_1)$ is conditionally given ξ_1 distributed according to Beta$(2, \frac{1.4 - 0.8 \cdot \xi_1}{0.3 + 0.4 \cdot \xi_1})$. A sample of 5000 points distributed according to P is shown in Figure 4.1.

For a given integers m_1 and m_2 we define the sets $A_{i,j}$ as the squares with vertices $[((i-1)/m_1, (j-1)/m_2); (i/m_1, (j-1)/m_2); ((i-1)/m_1, j/m_2), (i/m_1, j/m_2)]$ for $i = 1, \ldots, m_1 + 1; j = 1, \ldots, m_2 + 1$. The upper approximation sits on the $(m_1 + 1) \cdot (m_2 + 1)$ points $(i/m_1, j/m_2)$. The lower approximation sits on some barycenters of the $A_{i,j}$. The probabilities are $p_{i,j} = P(A_{i,j})$.

If $A_{i,j} = [(a, c); (a, d); (b, c); (b, d)]$ is such a rectangle, and (u_1, u_2) is a point in this rectangle, then let $P_{i,j}(u_1, u_2)$ be a probability measure sitting on the vertices with probability

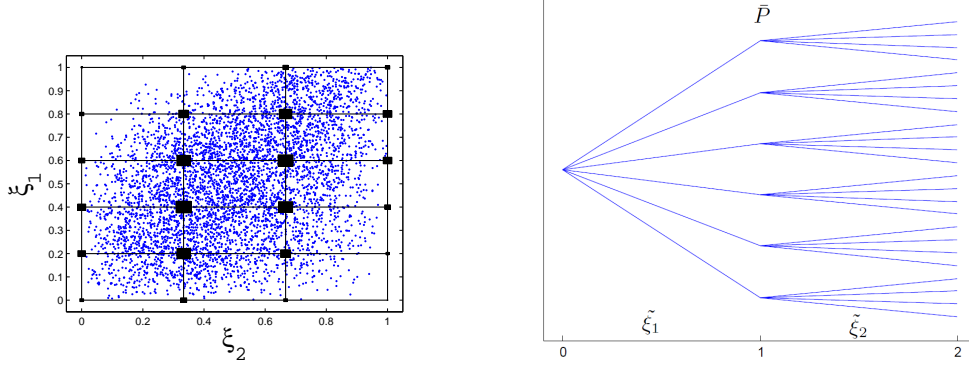$$p(a, c) = \frac{b - u_1}{b - a} \cdot \frac{d - u_2}{d - c}$$

FIG. 4.2. *The upper approximation $\bar{P}$ based on convex stochastic dominance (left) and the corresponding scenario tree structure (right).*

$$p(a, d) = \frac{b - u_1}{b - a} \cdot \frac{u_2 - c}{d - c}$$

$$p(b, c) = \frac{u_1 - a}{b - a} \cdot \frac{d - u_2}{d - c}$$

$$p(b, d) = \frac{u_1 - a}{b - a} \cdot \frac{u_2 - c}{d - c}.$$

Notice that the expectation of $P_{i,j} u_1, u_2$ is (u_1, u_2).

In order to estimate the upper and lower approximations $\bar{P}$ resp. $\underline{P}$, we use a large sample of N random deviates $(\xi_1^{(n)}, \xi_2^{(n)})$. Set

$$\bar{P} = \frac{1}{N} \sum_{n=1}^{N} P_{i,j}(\xi_1^{(n)}, \xi_2^{(n)}) \cdot \mathbb{1}_{(\xi_1^{(n)}, \xi_2^{(n)}) \in A_{i,j}}.$$

$\bar{P}$ is a finite process, and defines a tree process, which is the upper approximation. Figure 4.2 shows a construction of an upper approximation $\bar{P}$ based on convex stochastic dominance with the corresponding scenario tree structure with $m_1 = 5$ and $m_2 = 3$.

For the lower bound, the generation algorithm is a little more complicated. We sample $(\xi_1^{(n)}, \xi_2^{(n)})$, $n = 1, \ldots, N$ from P and set

$$n_{i,j} = \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{(\xi_1^{(n)}, \xi_2^{(n)}) \in A_{i,j}};$$

$$p_{i,j} = n_{i,j}/N;$$

$$z_1^{i,j} = \frac{1}{n_{i,j}} \sum_{n=1}^{N} \xi_1^{(n)} \mathbb{1}_{(\xi_1^{(n)}, \xi_2^{(n)}) \in A_{i,j}};$$

$$z_2^{i,j} = \frac{1}{n_{i,j}} \sum_{n=1}^{N} \xi_2^{(n)} \mathbb{1}_{(\xi_1^{(n)}, \xi_2^{(n)}) \in A_{i,j}};$$
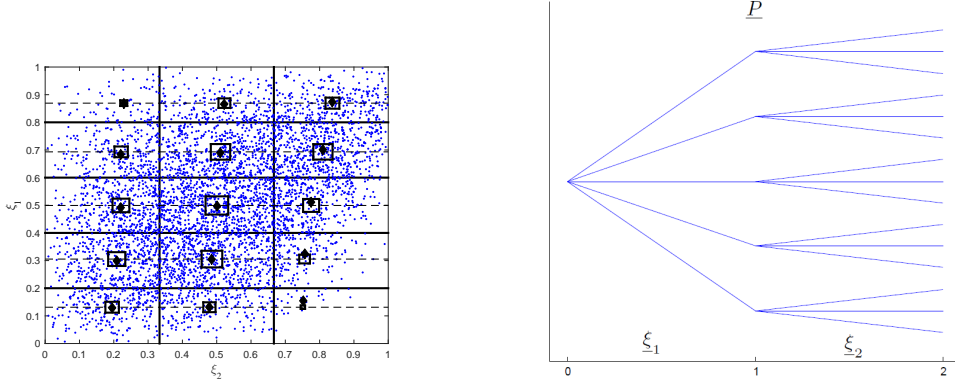
FIG. 4.3. *The barycenters (black diamonds) and the modified barycenters (black squares) of the lower approximation tree based on convex stochastic dominance (left) and the corresponding scenario tree structure (right).*

FIG. 4.4. *The upper approximation $\tilde{P}$ based on first order stochastic dominance (left) and the corresponding scenario tree structure (right).*

$$\bar{z}_1^i = \frac{\sum_j p_{i,j} z_1^{i,j}}{\sum_j p_{i,j}}.$$

Then $\underline{P}$ is defined as

$$\underline{P} = \sum_{i,j} p_{i,j} \delta_{(z_1^i, z_2^{i,j})}.$$

$\bar{P}$ can be represented as a tree. Arcs (i,j) for which $p_{i,j} = 0$ can be eliminated. Figure 4.3 shows the barycenters $(z_1^{i,j}, z_2^{i,j})$ (black diamonds) and the modified barycenters $(\bar{z}_1^i, z_2^{i,j})$ (black squares) for the choice $m_1 = 5$, $m_2 = 3$ and the distribution as in Figure 4.1.

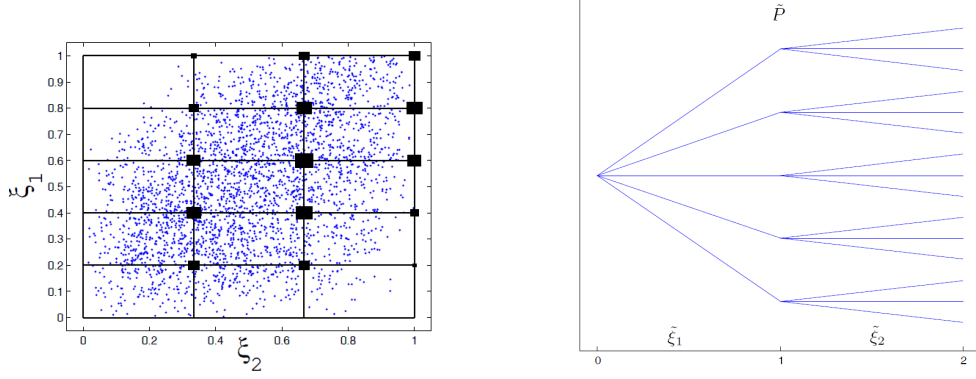Finally Figure 4.4 shows an upper approximation $\tilde{P}$ based on first order stochastic dominance with the corresponding scenario tree structure as described in Section 3.1. Similarly a lower approximation $\underset{\sim}{P}$ based on first-order stochastic dominance can be obtained by putting the weights to the left and lower corner of each rectangles in which the support has been dissected.

5. Case study: a multistage production problem. This section presents a simple multistage production problem adopted to test the bounds introduced before. The problem can be summarized as follows: Consider a single product inventory system, which is comprised of a warehouse and a factory. The planning horizon is T periods. Random demands have to be satisfied from an inventory (the only random quantities in the model). If the random demand exceeds the stock, it will be satisfied by rapid orders from a different source, which come at a higher price. At each time step (stage), orders can be placed. The goal is to minimize the total production cost of the factory in the entire planning period.

Let assume the following notation.

Deterministic parameters:

c_t the cost of producing a unit of the product
at the factory at time $t = 0, \ldots, T - 1$;

b_t the procurement cost from another retailer
for a unit of product at time $t = 1, \ldots, T$;

s_t the selling price at time $t = 1, \ldots, T$;

h_t the inventory holding costs for positive inventory
from time t to $t + 1$, $t = 0, \ldots, T - 1$;

d the final value of the inventory;

$\mathcal{P}_t$ the maximal production capacity
of factory at time $t = 0, \ldots, T - 1$;

v_0 is the amount of the product
in the warehouse at the beginning of the period 1;

Q the maximal cumulative production capacity
of the factory up to time $T - 1$.

Stochastic scenario process:

ξ_t is the demand for the product at time $t = 1, \ldots, T$.
(the random scenario process) All the demand must be satisfied;

$\boldsymbol{\xi}_t$ is the history of the demand for the product until time t.

Stochastic decision variables:

$x_t \geq 0$ is the amount of the product to be produced by the factory
and used to satisfy the demand at time $t = 0, \ldots, T - 1$.

Auxiliary variables:

v_t the amount of the product in the warehouse
after sales are effectuated at $t = 1, \ldots, T$.

Notice that if v_t is positive, $v_t = [v_t]_+$, an inventory holding cost $h_t \cdot [v_t]_+$ will be paid to carry the stock to the next step. If v_t is negative, $v_t = [v_t]_-$, a procurement cost $b_t [v_t]_-$ to buy extra stock from another retailer will be paid. The final stock is valuated with the value $d [v_T]_+$.

The problem can be modelled as follows:

$$\min \mathbb{E}[c_0 \cdot x_0 + h_0 \cdot v_0 + \sum_{t=1}^{T-1} c_t \cdot x_t(\boldsymbol{\xi}_t) + \sum_{t=1}^{T-1} h_t \cdot [v_t(\boldsymbol{\xi}_t)]_+ + \sum_{t=1}^{T} b_t \cdot [v_t(\boldsymbol{\xi}_t)]_-$$

$$-\sum_{t=1}^{T} s_t \cdot \xi_t - d \cdot [v_t(\boldsymbol{\xi}_T)]_+], \tag{5.1}$$

$$\text{s.t.} \quad 0 \leq x_0 \leq \mathcal{P}_0, \tag{5.2}$$

$$0 \leq x_t(\boldsymbol{\xi}_t) \leq \mathcal{P}_t \qquad t = 1, \ldots, T-1, \tag{5.3}$$

$$Q \geq x_0 + \sum_{t=1}^{T-1} x_t(\boldsymbol{\xi}_t), \tag{5.4}$$

$$v_1(\xi_1) = v_0 + x_0 - \xi_1, \tag{5.5}$$

$$v_{t+1}(\boldsymbol{\xi}_{t+1}) = [v_t(\boldsymbol{\xi}_t)]_+ + x_t(\boldsymbol{\xi}_t) - \xi_{t+1} \qquad t = 1, \ldots, T-1, \tag{5.6}$$

$$v_t(\boldsymbol{\xi}_t) = [v_t(\boldsymbol{\xi}_t)]_+ - [v_t(\boldsymbol{\xi}_t)]_- \qquad t = 1, \ldots, T, \tag{5.7}$$

$$[v_t(\boldsymbol{\xi}_t)]_+ \geq 0 \qquad t = 1, \ldots, T, \tag{5.8}$$

$$[v_t(\boldsymbol{\xi}_t)]_- \geq 0 \qquad t = 1, \ldots, T. \tag{5.9}$$

The objective function (5.1) denotes the expected total cost obtained from production, procurement from external retailers, inventory holding while the last two terms are the profits respectively from selling and for the final value of the inventory. Constraints (5.2)-(5.3) impose lower and upper levels on the factory production, constraint (5.4) imposes an upper bound on the maximal cumulative production capacity of the factory throughout the planning horizon. Finally constraints (5.5),(5.6), (5.7), (5.8) and (5.9) respectively define the dynamics of the inventory level and its definition.

5.1. Risk aversion strategy: including the Average Value at Risk. Given the confidence level α, we introduce now in the model (5.1)-(5.9), the (upper) average value at risk:

$$AVaR_\alpha = \min \{y + \frac{1}{(1-\alpha)} \mathbb{E}([c_0 \cdot x_0 + h_0 \cdot v_0 + \sum_{t=1}^{T-1} c_t \cdot x_t(\boldsymbol{\xi}_t)$$

$$+ \sum_{t=1}^{T-1} h_t \cdot [v_t(\boldsymbol{\xi}_t)]_+ + \sum_{t=1}^{T} b_t \cdot [v_t(\boldsymbol{\xi}_t)]_- - \sum_{t=1}^{T} s_t \cdot \xi_t$$

$$-d \cdot [v_t(\boldsymbol{\xi}_T)]_+ - y]_+) : y \in \mathbb{R}\}, \tag{5.10}$$

where y represents the Value at Risk (VaR). If $\alpha = 0$, then $AVaR_0$ equals the expectation and if $\alpha = 1$, then $AVaR_1$ is consistently defined as the essential supremum.

Introducing the auxiliary variable $u(\boldsymbol{\xi}_T)$ the model (5.1)-(5.9) in a risk aversion strategy becomes:

$$\min y + \frac{1}{(1-\alpha)} \mathbb{E} u(\boldsymbol{\xi}_T), \tag{5.11}$$

$$\text{s.t.} \quad u(\boldsymbol{\xi}_T) \geq c_0 \cdot x_0 + h_0 \cdot v_0 + \sum_{t=1}^{T-1} c_t \cdot x_t(\boldsymbol{\xi}_t)$$

$$+ \sum_{t=1}^{T-1} h_t \cdot [v_t(\boldsymbol{\xi}_t)]_+ + \sum_{t=1}^{T} b_t \cdot [v_t(\boldsymbol{\xi}_t)]_- - \sum_{t=1}^{T} s_t \cdot \xi_t$$

$$-d \cdot [v_t(\boldsymbol{\xi}_T)]_+ - y, \tag{5.12}$$

$$0 \leq x_0 \leq \mathcal{P}_0, \tag{5.13}$$

$$0 \leq x_t(\boldsymbol{\xi}_t) \leq \mathcal{P}_t \qquad t = 1, \dots, T-1, \tag{5.14}$$

$$Q \geq x_0 + \sum_{t=1}^{T-1} x_t(\boldsymbol{\xi}_t), \tag{5.15}$$

$$v_1(\xi_1) = v_0 + x_0 - \xi_1, \tag{5.16}$$

$$v_{t+1}(\boldsymbol{\xi}_{t+1}) = [v_t(\boldsymbol{\xi}_t)]_+ + x_t(\boldsymbol{\xi}_t) - \xi_{t+1} \qquad t = 1, \dots, T-1, \tag{5.17}$$

$$v_t(\boldsymbol{\xi}_t) = [v_t(\boldsymbol{\xi}_t)]_+ - [v_t(\boldsymbol{\xi}_t)]_- \qquad t = 1, \dots, T, \tag{5.18}$$

$$[v_t(\boldsymbol{\xi}_t)]_+ \geq 0 \qquad t = 1, \dots, T, \tag{5.19}$$

$$[v_t(\boldsymbol{\xi}_t)]_- \geq 0 \qquad t = 1, \dots, T, \tag{5.20}$$

$$u(\boldsymbol{\xi}_T) \geq 0. \tag{5.21}$$

5.2. Computation of bounds for a multistage risk-averse production problem. This section presents some computational tests on the three-stage $(T = 2)$ risk-averse production problem. We assume that the distribution P of the demand scenario process is given by $\boldsymbol{\xi}_2 = (\xi_1, \xi_2)$, where $\xi_1 \sim P_1$ is distributed according to a Beta(2,2) distribution and $\xi_2|\xi_1 \sim P_2(\cdot|\xi_1)$ is conditionally given ξ_1 distributed according to Beta$(2,(1.4 - 0.8 \cdot \xi_1)/(0.3 + 0.4 \cdot \xi_1))$ in the range [0,100]. The maximal production capacity of the factory at each period $t = 0,1$ is $\mathcal{P}_t = 567$ units, and the integral production capacity of the factory for the entire planning period is $Q = 13600$. The initial inventory is $v_0 = 10$, the final value of the inventory is $d = 2$ per unit, and values of production price c_t, selling price s_t, inventory holding cost h_t and procurement cost b_t at time period t are presented in Table 5.1.

TABLE 5.1
Production price c_t, selling price s_t, inventory holding cost h_t from time t to time $t+1$ and procurement cost b_t for extra stock from another retailer at time t.

t	c_t	s_t	h_t	b_t
0	3.5	-	2	-
1	3.6	10.7	1.9	8
2	-	10.5	-	8.1

The linear problems derived from our case study have been solved under *AMPL* environment along *MOSEK* solver by interior-point algorithm. All the computations have been performed on a 64-bit machine with 12 GB of RAM and a 2.90 GHz processor.

In order to find guaranteed bounds, we consider first two finite three-stage trees $\tilde{P} = (\tilde{\xi}_1, \tilde{\xi}_2)$ and $\underline{P} = (\underline{\xi}_1, \underline{\xi}_2)$ having the same structure $\text{T}_{5,5}$: 5 branches from the root and 5 from each of the second-stage nodes resulting in $k = 5 \times 5 = 25$ scenarios and 31 nodes. The two finite scenario trees have been built according to first order stochastic dominance as described in Section 3.1 providing respectively upper and

lower bounds: They are obtained by dissecting the support intro 25 rectangles $A_{i,j}$, $i = 1, \ldots, 5$, $j = 1, \ldots, 5$ and putting the weights respectively to the left and lower corner (a_i, b_j) and to the up and right corner (a_{i+1}, b_{j+1}). Similarly other pairs of finite scenario trees with bushier tree structures $T_{10,10}, T_{20,20}, T_{40,40}, T_{80,80}$ and $T_{160,160}$ have been considered (see Table 5.2 for details). Lower and upper bounds to the total cost of problem (5.11)-(5.21) by using the finite scenario trees based on first order stochastic dominance (FSD) are reported in Tables 5.3-5.4-5.5 and Figures 5.1-5.2. As expected the worst lower and upper bounds are given by $T_{5,5}$ with an absolute gap $v(\tilde{P}) - v(\underline{P})$ of 137.5 but requiring the lowest CPU time (0.0625 CPU seconds over 30 runs). Increasing the size of the scenario tree, signicantly improves the bounds, monotonically reaching lower values of gaps up to 4.11 for the biggest scenario tree considered $T_{160,160}$ (see Figure 5.1 in the case of $\alpha = 0$ where the bounds are plotted for increasing values of complexity of calculation measured in CPU seconds). Similar results are obtained for different values of confidence level α (see Figure 5.2). The time required to solve the problem (see last column of Tables 5.3-5.4), monotonically increases with the dimension of the tree reaching the highest value for $T_{160,160}$ (3.40625 CPU seconds over 30 runs). Finally, average relative gaps $\frac{v(\tilde{P}) - v(\underline{P})}{v(\underline{P})}$ are reported in table 5.5: as expected they improve monotonically with the number of scenarios in the trees, ranging from 32% for $T_{5,5}$ to 0.8% for $T_{160,160}$.

TABLE 5.2
Scenario tree structures based on first order stochastic dominance (FSD) and convex order dominance (CXD) adopted to compute the bounds.

	Tree	Number of scenarios	Number of nodes
FSD - CXD	$T_{5,5}$	25	31
FSD - CXD	$T_{10,10}$	100	111
FSD - CXD	$T_{20,20}$	400	421
FSD	$T_{40,40}$	1600	1641
FSD	$T_{80,80}$	6400	6481
FSD	$T_{160,160}$	25600	25761

TABLE 5.3
Lower bounds objective function values and complexity of calculation (in CPS seconds) of finite scenario tree structures, based on first order stochastic dominance (FSD) for increasing values of α.

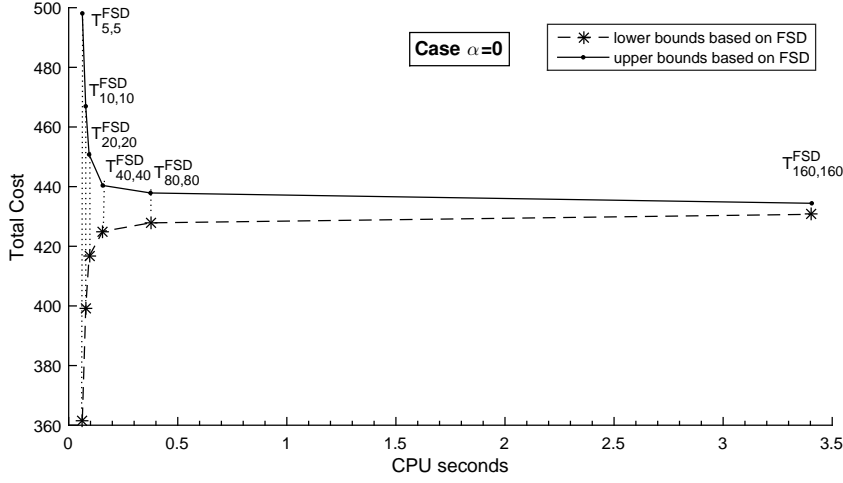| Trees	$v(\underline{P})$						CPU seconds
	$\alpha = 0$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.5$	$\alpha = 0.7$	$\alpha = 0.9$	
$T_{5,5}^{FSD}$	361.66	379.61	414.36	450.22	479.13	514.65	0.062
$T_{10,10}^{FSD}$	399.09	415.49	447.27	479.49	514.06	565.42	0.078
$T_{20,20}^{FSD}$	416.66	433.89	464.34	495.87	531.56	582.08	0.093
$T_{40,40}^{FSD}$	424.98	441.91	472.85	503.94	539.45	589.18	0.156
$T_{80,80}^{FSD}$	427.86	445.24	477.25	509.58	545.95	593.42	0.375
$T_{160,160}^{FSD}$	430.71	448.02	479.51	511.47	547.34	596.37	3.406

Fig. 5.1. *Lower and upper bounds to the total cost of problem (5.11)-(5.21) with confidence level $\alpha = 0$, by using the finite scenario trees based on first order stochastic dominance, for increasing values of complexity of calculation measured in CPU seconds.*
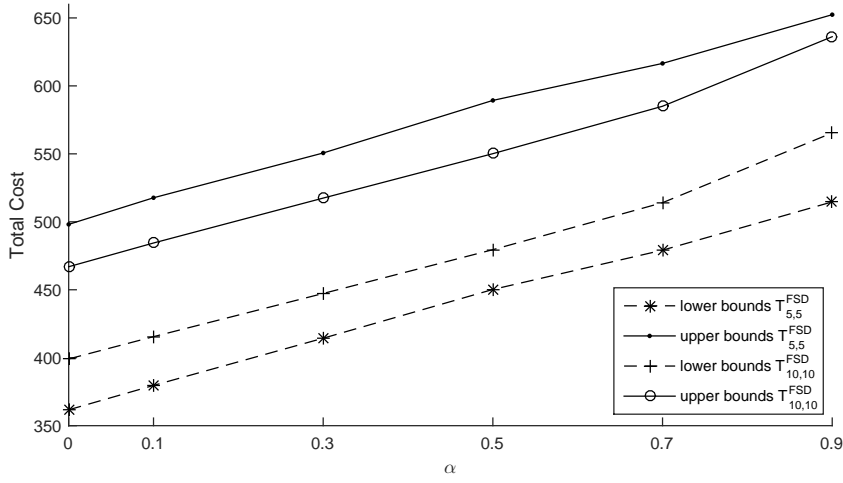

Fig. 5.2. *Lower and upper bounds to the total cost of problem (5.11)-(5.21) by using the finite scenario trees $T_{5,5}$ and $T_{10,10}$ based on first order stochastic dominance.*

We consider now lower and upper bounds built on convex stochastic dominance as described in Section 3.2. They are obtained by dissecting the support into $m_a \times m_b$ rectangles $A_{i,j}$, $i = 1, \ldots, m_a$, $j = 1, \ldots, m_b$ and putting the weights respectively to the barycenter and to the four corners. In this way the bounds can be calculated on two finite trees without evaluating the continuous problem.

Lower and upper bounds based on convex stochastic dominance (CXD) are reported in Tables 5.6-5.7-5.8 and Figures 5.3-5.4. Since ξ_2 depends by ξ_1, the correction term described in Section 3.2.2 for lower tree approximation should be computed (see

Table 5.6, column 3). This is obtained as follows: problem (5.1)-(5.9) can be rewritten as

$$\min \mathbb{E}[c_0 \cdot x_0 + h_0 \cdot v_0 + \sum_{t=1}^{T-1} c_t \cdot x_t(\boldsymbol{\xi}_t) + \sum_{t=1}^{T-1} h_t \cdot [v_t(\boldsymbol{\xi}_t)]_+ + \sum_{t=1}^{T} b_t \cdot [v_t(\boldsymbol{\xi}_t)]_-$$

$$- \sum_{t=1}^{T} s_t \cdot \xi_t - d \cdot [v_t(\boldsymbol{\xi}_T)]_+ + \Psi[x_0, \dots, x_T, \boldsymbol{\xi}_T]],$$

where

$$\Psi[x_0, \dots, x_T, \boldsymbol{\xi}_T] = \begin{cases} 0 & \text{if } (x_0, \dots, x_T) \in \mathbb{X} \\ \infty & \text{otherwise} \end{cases}$$

with

$$\mathbb{X} := \begin{cases} 0 & \leq & x_0 \leq \mathcal{P}_0, \\ 0 & \leq & x_t(\boldsymbol{\xi}_t) \leq \mathcal{P}_t & t = 1, \dots, T-1, \\ Q & \geq & x_0 + \sum_{t=1}^{T-1} x_t(\boldsymbol{\xi}_t), \\ v_1(\xi_1) & = & v_0 + x_0 - \xi_1, \\ v_{t+1}(\boldsymbol{\xi}_{t+1}) & = & [v_t(\boldsymbol{\xi}_t)]_+ + x_t(\boldsymbol{\xi}_t) - \xi_{t+1} & t = 1, \dots, T-1, \\ v_t(\boldsymbol{\xi}_t) & = & [v_t(\boldsymbol{\xi}_t)]_+ - [v_t(\boldsymbol{\xi}_t)]_- & t = 1, \dots, T, \\ [v_t(\boldsymbol{\xi}_t)]_+ & \geq & 0 & t = 1, \dots, T, \\ [v_t(\boldsymbol{\xi}_t)]_- & \geq & 0 & t = 1, \dots, T. \end{cases}$$

Let $\underline{P}$ be the tree constructed using $\tilde{z}_1^i$ as first stage values and and $z_2^{i,j}$ as second stage values with scenario probabilities $p_{i,j}$. According to Proposition 3.5, we have that the error made by the tree process $\underline{P}$ for our three-stage production problem by collapsing $z_1^{i,j}$ in $\tilde{z}_1^i$, $i = 1, \dots, m_a$ is $c_1 \sum_{i,j} p_{i,j} \cdot |\tilde{z}_1^i - z_1^{i,j}|$. In the risk-adverse case we need just to divide the previous expression by the confidence level $(1 - \alpha)$. Notice that if the demand at time 2 is independent by the demand at time 1 then the error is null. The absolute gap between CXD lower and upper bounds based on the simplest tree structure considered $T_{5,5}$, reduces considerably compared to the one obtained by first order construction passing, in case of $\alpha = 0$, from to 136.44 to 11.5 units. Increasing the size of the scenario tree to $T_{20,20}$, signicantly improves the bounds closing the gap (see Figure 5.4) instead of 34.1 for FSD and taking approximately the same CPU time. Different values of confidence level α are considered in Figure 5.3 and relative gaps $\frac{v(\bar{P}) - v(\underline{P})}{v(\underline{P})}$ are reported in table 5.8: results show that the average gaps considerably reduces passing from 4% with $T_{5,5}^{CXD}$ to 0% with $T_{20,20}^{CXD}$.

6. Conclusions. The paper develops lower and upper bounds for multistage stochastic programs based on first order stochastic dominance and convex order dominance of probability measures. The proposed method allows to construct solutions for the infinite problem by considering finite trees approximations as proxies, and can even be made arbitrarily close by making the approximating trees bushier. For illustration numerical results on a multistage risk-averse production problem are presented. Results show that the solutions based on convex order dominance construction outperform the ones obtained by first order stochastic dominance closing the gap between upper and lower bounds within a limited computational complexity and simple scenario tree structures.
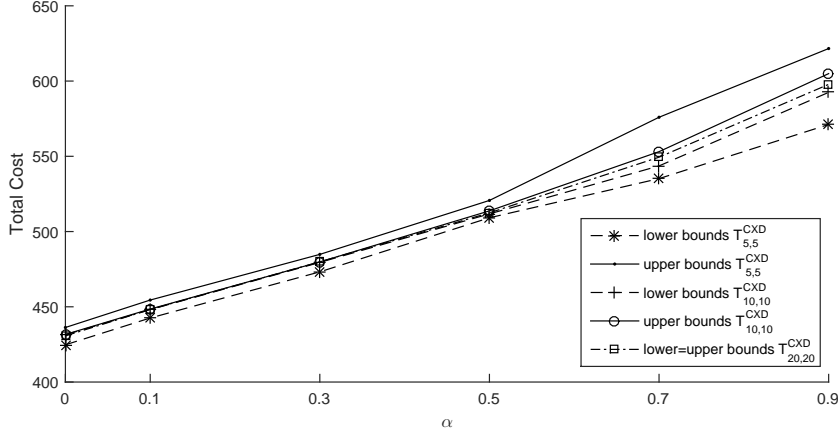
FIG. 5.3. *Lower and upper bounds to the total cost of problem (5.11)-(5.21) by using the finite scenario trees $T_{5,5}$, $T_{10,10}$ and $T_{20,20}$ based on convex stochastic dominance (CXD).*
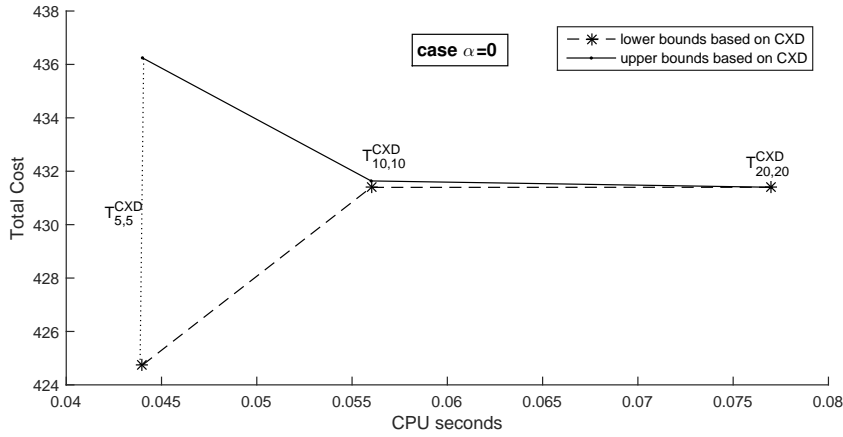

FIG. 5.4. *Lower and upper bounds to the total cost of problem (5.11)-(5.21) with $\alpha = 0$, by using finite scenario trees based on convex dominance (CXD).*

REFERENCES

[1] J. R. BIRGE, *The value of the stochastic solution in stochastic linear programs with fixed recourse*, Math. Program., 24 (1982), pp. 314–325.

[2] N. BOLAND, I. BAKIR, B. DANDURAND AND A. ERERA, *Scenario Set Partition Dual Bounds for Multistage Stochastic Programming: A Hierarchy of Bounds and a Partition Sampling*, www.optimization-online.org, 2016.

[3] L.F. ESCUDERO, A. GARÍN, M. MERINO AND G. PÉREZ, *The value of the stochastic solution in multistage problems*, TOP, 15 (2007), pp. 48–64.

[4] K. FRAUENDORFER AND M. SCHÜRLE, *Multistage stochastic programming: Barycentric approximation*, P. Pardalos, and C. A. Floudas eds., Encyclopedia of Optimization, Vol. 3 Dordrecht, NL: Kluwer Academic Publishers, ISBN 0-7923-6932-7, 2001, pp. 576–580.

[5] K. FRAUENDORFER, D. KUHN AND M. SCHÜRLE, *Barycentric Bounds in Stochastic Programming: Theory and Application*, In Stochastic Programming: The State of the Art, In Honor of George B. Dantzig, Gerd Infanger ed., Springer, New York, 2011, pp. 67–96.

[6] C.C. HUANG, I. VERTINSKY AND W.T. ZIEMBA, *On multiperiod stochastic dominance*, J. Fi-

nanc. Quant. Anal., 13(1) (1978), pp. 1–13.

[7] J.L. JENSEN, *Sur les fonctions convexes et les inégalités entre les valeurs moyennes*, Acta Mathematica, 30(1) (1906), pp. 175–193.

[8] D. KUHN, *Generalized Bounds for Convex Multistage Stochastic Programs*, Lecture Notes in Economics and Mathematical Systems 548, Springer-Verlag, Berlin, 2005.

[9] D. KUHN, *Aggregation and discretization in multistage stochastic programming*, Math. Program., Ser. A, 113 (2008), pp. 61–94.

[10] H. LEVY AND J. PAROUSH, *Multi-Period Stochastic Dominance*, Manag. Sci., 21 (1974), pp. 428–435.

[11] A. MADANSKY, *Bounds on the expectation of a convex function of a multivariate random variable*, Ann. Math. Stat., 30(3) (1959), pp. 743–746.

[12] A. MADANSKY, *Inequalities for stochastic linear programming problems*, Manag. Sci., 6 (1960), pp. 197–204.

[13] F. MAGGIONI, E. ALLEVI AND M. BERTOCCHI, *Bounds in Multistage Linear Stochastic Programming*, J. Optim. Theory Appl., 163(1) (2014), pp. 200–229.

[14] F. MAGGIONI, E. ALLEVI AND M. BERTOCCHI, *Monotonic bounds in multistage mixed-integer linear stochastic programming*, Comput. Manag. Sci., 13(3) (2016), pp. 423–457.

[15] F. MAGGIONI AND G. CH. PFLUG, *Bounds and approximations for multistage stochastic programs*, Siam J. Optim., 26(1) (2016), pp. 831–855.

[16] G. MOKOBODSKI, *Elements extrémeaux pour le balayage*, Séminare Brelot-Choquet-Deny, 13-ième année, 5 (1970).

[17] A. MÜLLER ADN D. STOYAN, *Comparison Methods for Stochastic Models and Risks*, Wiley, 2002. (2002).

[18] G. CH. PFLUG, *Version-independence and nested distributions in multistage stochastic optimization*, SIAM J. Optim., 20(3) (2009), pp. 1406–1420.

[19] G. CH. PFLUG AND A. PICHLER, *Multistage Stochastic Programming*, (Springer Series in Operations Research and Financial Engineering), Springer, 2014.

[20] B. SANDIKÇI, N. KONG AND A.J. SCHAEFER, *A hierarchy of bounds for stochastic mixed-integer programs*, Math. Program., Ser. A, 138(1) (2012), pp. 253–272.

[21] B. SANDIKÇI AND O.Y. ÖZALTIN, *A scalable bounding method for multistage stochastic integer programs*, www.optimization-online.org, 2014.

[22] G.L. ZENAROSA, O.A. PROKOPYEV AND A.J. SCHAEFER, *Scenario-Tree Decomposition: Bounds for Multistage Stochastic Mixed-Integer Programs*, www.optimization-online.org, 2014.

TABLE 5.4

Upper bounds objective function values and complexity of calculation (in CPS seconds) of finite scenario tree structures, based on first order stochastic dominance (FSD) for increasing values of α.

Trees	$v(\tilde{P})$						CPU seconds
	$\alpha = 0$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.5$	$\alpha = 0.7$	$\alpha = 0.9$	
$T_{5,5}^{FSD}$	498.11	517.68	550.63	589.31	616.53	652.39	0.062
$T_{10,10}^{FSD}$	466.95	484.41	517.45	550.13	585.02	636.19	0.078
$T_{20,20}^{FSD}$	450.76	468.31	498.78	530.31	566.00	616.52	0.093
$T_{40,40}^{FSD}$	440.40	457.90	490.28	522.79	558.68	606.50	0.156
$T_{80,80}^{FSD}$	437.88	455.18	486.54	518.54	554.38	604.34	0.375
$T_{160,160}^{FSD}$	434.42	451.76	483.42	515.61	551.73	601.18	3.406

TABLE 5.5

Gaps of finite scenario tree structures based on first order stochastic dominance (FSD) for increasing values of α.

Trees	$\frac{v(\tilde{P})-v(\underline{P})}{v(\underline{P})}$					
	$\alpha = 0$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.5$	$\alpha = 0.7$	$\alpha = 0.9$
$T_{5,5}^{FSD}$	0.377	0.363	0.328	0.308	0.286	0.267
$T_{10,10}^{FSD}$	0.170	0.165	0.156	0.147	0.138	0.125
$T_{20,20}^{FSD}$	0.081	0.079	0.074	0.069	0.064	0.059
$T_{40,40}^{FSD}$	0.036	0.036	0.036	0.037	0.035	0.029
$T_{80,80}^{FSD}$	0.023	0.022	0.019	0.017	0.015	0.018
$T_{160,160}^{FSD}$	0.008	0.008	0.008	0.008	0.008	0.008

TABLE 5.6

Lower bounds objective function values, errors $c_1 \sum_{i,j} p_{i,j} \cdot |\tilde{z}_1^i - z_1^{i,j}|$ and complexity of calculation (in CPS seconds) of finite scenario tree structures, based on convex stochastic dominance (CXD) for increasing values of α.

Trees	$v(\underline{P})$						Error	CPU s.
	$\alpha = 0$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.5$	$\alpha = 0.7$	$\alpha = 0.9$		
$T_{5,5}^{CXD}$	424.74	442.55	473.21	509.15	535.18	571.15	0.0044	0.046
$T_{10,10}^{CXD}$	431.40	448.12	479.49	511.74	543.62	592.49	0.0031	0.059
$T_{20,20}^{CXD}$	431.40	448.21	479.81	512.43	549.41	597.81	0.0027	0.076

TABLE 5.7
Upper bounds objective function values and complexity of calculation (in CPS seconds) of finite scenario tree structures, based on convex stochastic dominance (CXD) for increasing values of α.

Trees	$v(\bar{P})$						CPU seconds
	$\alpha = 0$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.5$	$\alpha = 0.7$	$\alpha = 0.9$	
$T_{5,5}^{CXD}$	436.25	454.48	484.83	520.67	575.95	621.62	0.042
$T_{10,10}^{CXD}$	431.64	448.50	479.82	513.74	553.09	604.95	0.054
$T_{20,20}^{CXD}$	431.40	448.21	479.81	512.43	549.41	597.81	0.078

TABLE 5.8
Gaps of finite scenario tree structures based on convex dominance (CXD) for increasing values of α.

Trees	$\frac{v(\bar{P}) - v(\underline{P})}{v(\underline{P})}$					
	$\alpha = 0$	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.5$	$\alpha = 0.7$	$\alpha = 0.9$
$T_{5,5}^{CXD}$	0.027	0.026	0.024	0.022	0.076	0.088
$T_{10,10}^{CXD}$	0.0005	0.0008	0.0006	0.003	0.017	0.020
$T_{20,20}^{CXD}$	0	0	0	0	0	0